Machine Learning for Sample Selection Models: A First Report

Owen McDevitt[1]

***Abstract*—**

- **Guidelines for using ML and sample selection.**
- **Simulations, model, selection and outcome**
- **Compare bias/efficinecy of methods.**
- **Results**
- **Implications**

I. Introduction

- State of literature (very little)
- Maybe background on problem of sample selection?
- Maybe background on machine learning methods to be used in paper?
- Honestly, I am not exactly sure how to format the paper and what to include/exclude.

II. Model/Simulations

(1) $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ (2) $U_i = \gamma_0 + \gamma_1 Z_i + \eta_i$

- Model for simulations
- Justification for the methods (math?)
- Show mathematically how the machine learning corrects the selection bias

III. Results/Applications

- what methods worked best (error), were most efficient (computation time), etc.
- Potential applications of the results

IV. CONCLUSIONS

Sorry it's such a rough, rough draft. I will work on it more once I finish the code.

References

[1] Chernozhukov papers?
[2] Millimet Tschernis JBES paper?
[3] Other suggestions?

[1]Owen McDevitt is a student at the University of Oklahoma majoring in Mathematics and Economics, ocmdevitt1022@ou.edu